DOCUMENT RESUME

ED 249 767 FL 014 540

AUTHOR Ingram, D. E.
TITLE Using Proficiency Rating Scales with High School Foreign Language Learners.
PUB DATE 84
NOTE 18p.; Paper presented at the Annual Meeting of the Australian Federation of Modern Language Teachers Associations (5th, Hobart, Australia, August 31-September 4, 1984).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Curriculum Development; Foreign Countries; High School Students; *Language Proficiency; *Language Tests; Measurement Techniques; Program Development; *Rating Scales; Secondary Education; *Second Language Instruction; *Test Use
IDENTIFIERS Australia

ABSTRACT

The use of language proficiency rating scales in Australian school language programs is discussed, focusing primarily on the Australian Second Language Proficiency Ratings (ASLPR), the only direct instrument developed in Australia that is in widespread use and that has been extensively tested. It is also the only one developed for use with a variety of languages (including English as a second language) and a variety of learner types, and also for measuring both general and special purpose proficiency. The appropriateness of the ASLPR at different program stages, the purposes for testing and the considerationr in matching purposes with instruments, item selection, and the use of both rating scales and other direct and semidirect test instruments in program planning are examined. The particular problems of relating ratings to grades are considered, with special attention given to the use of goals and objectives in the development of general language proficiency and to their implications for testing. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED249767

USING PROFICIENCY RATING SCALES WITH HIGH SCHOOL FOREIGN LANGUAGE LEARNERS

D.E. INGRAM

DEPARTMENT OF EDUCATION STUDIES,

DARWIN COMMUNITY COLLEGE.

Paper presented at the Fifth National Languages Conference of the Australian Federation of Modern Language Teachers Associations, Rosny College, Hobart, Tasmania, 31st August - 4th September, 1984.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

D. E. Ingram

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

FL 014 540

USING PROFICIENCY RATING SCALES

WITH HIGH SCHOOL FOREIGN LANGUAGE LEARNERS

ABSTRACT

In most States there is interest in more effective ways of stating learners' language proficiency and in how such statements might relate to graded objectives. This paper will consider how proficiency rating scales (especially the Australian Second Language Proficiency Ratings) can be used in school language programmes, their appropriateness at different stages, the purposes for which one tests and the instruments most suitable for them, and the use of test instruments (especially rating scales but also indirect and semi-direct tests) in programme planning. The particular problems posed by graded objectives courses will be considered with special attention to how they relate to the development of general proficiency and to the implications for testing.

CONTENTS

I Preamble
I.1 Scope of the Paper
I.2 Assessment, Testing and Tests
II Proficiency and Rating Scales
III Proficiency Rating Scales in School Assessment
- III.1 When to use
- III.2 How to use
- III.3 Continuous vs Exit - Point Assessment
- III.4 Selecting Items
- III.5 Relating Ratings and Grades through Goals

IV Tests in Programme Planning
V Graded Objectives
VI Conclusion

I Preamble : I.1 Scope of the Paper

In the two years since the last National Languages Conference, there has been growing interest in Australian Secondary Schools in the use of proficiency rating scales to assess foreign language achievement. In South Australia this has taken the form, in particular, of a "graded objectives" or "graded levels of achievement" approach. In New South Wales, the ACT, and Queensland (1), interest has been in using the Australian Second Language Proficiency Ratings (ASLPR) or some other "direct instrument" based on behavioural descriptions. Despite significant differences, these two approaches are similar in that they each seek to identify language proficiency by describing the sorts of tasks learners can carry out and how they are carried out.

In my paper to the 1982 conference, I discussed the ASLPR, the only direct instrument developed in Australia that is in widespread use both here and overseas and the only one extensively trialled and developed for use with a variety of languages (including ESL), with a variety of different learners (adolescents and adults, literate and illiterate), and in measuring both general and "special purpose" proficiencies (a current research project). The ASLPR is now so well-known that there is no need to describe it here (though its developers would be happy to provide any further information to inquirers). This paper will consider how the ASLPR or any other adequately developed and trialled behaviour-based scale can be used in school language programmes, its appropriateness at different stages and for different testing purposes, and the use of such scales and other test types in programme planning. Their relationship to graded objectives and the nature of assessment in a graded objectives course will also be considered.

I.2 Assessment, Testing and Tests

In this paper "assessment" or "testing" refers to the considered judgement that teachers make about student learning or student performance. "Test" refers to any activity used in evaluating or measuring some part or all of a learner's language proficiency or other performance [cf. Ingram 1984]. What tests or test types are used in the process of assessing language learning should depend on the aim of the assessment or testing. Thus, if one's interest is in how well the learner has formally learned the actual content presented during a course, then tests based specifically on that content, i.e., on the syntax, vocabulary or functions specified in the course content, will be most appropriate. If the aim of the testing is to measure progress in a "graded objectives" or "graded levels of achievement" course, then the test should focus on the situations, tasks, functions, and their exponents specified in the course content. If, however, the aim is to assess the learner's general proficiency (or "competency" as it is called in the new "ROSBA" syllabuses in Queensland), then a test that gives a picture of the learner's practical language skills in general, everyday situations will be most appropriate. Proficiency rating scales such as the ASLPR have been

developed and seem most suitable for this purpose, i.e., to measure general proficiency. However, they cannot be used continuously or indiscriminately in language programmes and the need is to realize what they are best used for, how they can best be used, and when it is appropriate to use them.

II Proficiency and Rating Scales :

Proficiency can be conceived of as the ability to carry out language tasks. In the ASLPR and other rating scales, proficiency levels are defined by describing the sorts of language tasks learners can carry out and how they are carried out in terms of such features as sentence forms, vocabulary range, formal accuracy, pronunciation, fluency, comprehensibility, rate of utterance, register flexibility, and so on. Whereas rating scales designed to measure general proficiency describe the tasks and how they are carried out in general terms, in terms of an overall picture of the language behaviour observable at each level as learners attempt to use the language to carry out the tasks, in the graded objectives or graded levels of achievement approach, the actual tasks and the actual exponents by which they are carried out are generally specified. In using rating scales to measure proficiency, one essentially asks what tasks the learners can carry out, how they carry them out and, most of all, what is the nature of the language behaviour observed as the learners try to communicate and what behavioural description on the scale best matches the observed behaviour.

One important advantage of rating scales such as the ASLPR over other test types is that they are developmental in structure, i.e., the progression of behavioural descriptions from zero to native-like proficiency attempts to reflect the way in which a second language develops from zero to native-like. Thus the scale is not arbitrary but related to the universal developmental schedule and a learner's rating indicates the stage of development he has reached in the developmental schedule. This is important for several reasons including at least three relevant to this paper. First, the proficiency measurements provided are not arbitrary in the sense that a percentage or a score such as 4 on a 7-point scale or 495 out of a possible 550 is arbitrary and meaningless but, in that they are behavioural and task-oriented, they tell the user exactly what the learner is capable of doing and, in that they relate to the developmental schedule, the behaviour specified exists within a progressive, developmental sequence of behaviours. Second, through the developmental schedule, there are changes in all aspects of language behaviour (syntax, vocabulary, discourse, functions, and so on) but nevertheless certain aspects of language behaviour become salient at different developmental stages and different parameters of change become salient at different stages. This is important both for the next reason to be discussed and because the key parameters help the rater to identify where the learner's proficiency falls on the scale. These parameters include, at the lowest levels, the range of situations in which the language can be used and a change from formulaic to more creative language, satisfaction of survival needs, emerging complexification that facilitates the creation of social

relationships (ASLPR 1+ onwards), register sensitivity and flexibility (3), and ability to see beyond surface meaning and to draw on significant cultural understanding (4 to 5). Third, because rating scales such as the ASLPR are behavioural and developmental, they indicate to the syllabus writer what learners at different levels can do, what their future directions of development are, and, therefore, what is most relevant to include in their course in order to maximize their rate of development.

III Proficiency Rating Scales in School Assessment :

III.1 When to Use

Rating scales such as the ASLPR are designed to measure learners' practical proficiency; therefore, if a course aims at developing learners' practical language skills, proficiency, ability to communicate, or, in the terms used in Queensland's new syllabuses, "competency", then a behavioural rating scale would seem the most appropriate instrument. More traditional grammar-translation or discrete-point tests that essentially test knowledge rather than proficiency would not seem appropriate [see Ingram 1984 for a fuller discussion]. However, general proficiency changes slowly especially in the unintensive courses offered in Secondary Schools and the best available figures suggest that it would take considerably more than 240 hours for the average Secondary School learner's proficiency to increase by one step on the ASLPR in, for example, French, German, or Italian [see FSI 1983]. Consequently, there is probably no point in using tests of general proficiency (including rating scales) any more frequently than on entry and exit or, at most, once per year. For on-going assessment through a course, the teacher's main interest is probably in the extent to which learners are assimilating the course content being presented and, for this purpose, more specific test types (both indirect and semi-direct) would seem appropriate.

III.2 How to Use

Most teachers would believe that, during a course, they develop a fairly accurate picture of the proficiency of each learner in their class. However, it would be highly unreliable for teachers simply to sit and rate each class member at the end of a course. First, they are likely to be unduly swayed by their knowledge of the course content, believing the students ought to be able to do certain things because they were taught. Second, they are likely to believe the extroverted child who is always contributing to the lesson and processes the language overtly to be more proficient than the introverted child or those who seem to process the language internally. Third, as is implied in the first two reasons, the rating, which is always subjective to allow for the complexity of language [cf. Ingram 1982 : Sections 3 and 4], is likely to become unduly impressionistic and to suffer in both validity and reliability. Validity and reliability may be increased if the scale is applied with correct procedures.

First, the teacher could keep a clipboard handy throughout the course in order to accumulate detailed checklists of the learners' behaviour based, for example, on the descriptions of proficiency levels contained in the ASLPR. This practice would reduce the teacher's reliance on her memory of each student's language behaviour and help to re-direct the focus of her attention from the course content (which the student might have formally learned yet not have available for use in real communication activities) to the behavioural descriptions in the scale which are designed to match practical proficiency and constrain intuitive judgements [cf. Ingram 1982]. Second, and most importantly, rating scales should be used in interviews structured for the purpose of rating. In principle, one could gain a picture of a learner's practical proficiency simply by observing him or her over a long period of time but, in a class, the teacher is unlikely to differentiate the learners adequately especially in the unintensive and formal classes of a Secondary School. In addition, there is a real danger that learners' actual proficiencies will not be observed because unstructured situations simply may not extend them fully. Consequently, interviews of the sort recommended for the ASLPR [see Ingram 1982 : Section IV; 1984 : Section III.3.ii] should always be used when learners are to be rated even though, for teachers, the interviews may be briefer and confirmatory. Such considerations raise the controversial issue of continuous versus exit-point (or end-of-course) assessment.

III.3 Continuous vs Exit-Point Assessment

It has become fashionable in education to speak favourably of continuous assessment and to deplore one-off end-of-course or exit-point assessment. However, such criticism of exit-point assessment ignores the nature of skill development in general as well as the distinctive nature of language learning. First, if a learner is to commence another course, start work, go overseas, make some other use of the language or simply leave the programme, it is his or her proficiency at that point that is of interest and not what it was 6, 12 or 24 months earlier. If, as suggested earlier, the teacher's continual monitoring of each learner's performance facilitates the final assessment, checklists of the sort discussed may be worthwhile but an exit-point test using the interview procedures recommended for use with rating scales is essential.

Second, there is an obvious and important distinction to be made between a long-term goal (such as attainment of a particular proficiency level) and the path followed in attaining that goal. In language learning, this is seen where different courses follow different paths yet aim at and develop the same levels of general proficiency. In addition, Krashen [e.g., 1976] has drawn attention to the fundamental distinction between "learning" and "acquisition", between what is formally taught, learned or memorized and what is able to be used in practical language tasks reflecting the creativity and spontaneity of real language use and general proficiency. Briefly, this means that a student may be able to formally learn aspects of the language (e.g., through rote memorization and the repetitive application of rules in formal exercises) and yet be unable to use the

rules or vocabulary when it comes to real-life communication tasks. If we are interested in measuring practical language skills or practical proficiency, it is the learners' proficiency as reflected in real-life communication tasks that is important and not their formal knowledge. Continuous assessment is more likely to focus on the latter (i.e., on their formal learning of the course content most recently taught) and is unlikely to provide a true picture of what has been acquired and is available for use spontaneously and creatively in practical language tasks.

For these reasons, rather than base proficiency or competency assessment on continuous testing, it would seem desirable for teachers to place most emphasis on exit-point assessment. What then is critical is that this summative assessment be so conducted that the learners' maximum language behaviour is elicited, observed and rated. This means that, for instance, interviews and other tests must be in accord with recommended principles [cf. Ingram 1982: Section IV; 1984: Section III.3.ii] and should be conducted, scheduled and organized so that each learner has an opportunity to demonstrate his or her real ability and is not disadvantaged by, for example, poor or unduly stressful procedures or such irrelevant issues as a day's illness.

Third, the contrast between continuous and exit-point assessment highlights again the fundamental distinction in testing between course attainment and general proficiency. Course attainment relates more to what learners have formally learned, to what they have taken from each of the teacher's lessons. Such testing can relate quite specifically to the content of the course and the test types could reflect quite closely the activities used in teaching. In fact, it is commonly "accepted wisdom" in education that assessment procedures should mirror the teaching and learning activities of the course. However, as demonstrated in the grammar-translation era of language teaching, tests of this sort, even though they may accurately reflect learners' attainments in the actual course taught, may give little information on the learners' real practical proficiency. Tests of course attainment may, in other words, be highly desirable for many purposes (e.g., feedback on the effectiveness of teaching strategies, success of formal teaching and learning, or for diagnostic purposes); they may, for these purposes, take many forms both indirect (e.g., discrete point tests) and semi-direct (e.g., cloze or dictation); and they will be used throughout a course. However, they should not be confused with measurements of general proficiency and, if the aim of the course is to develop general proficiency, the final assessment should be based around test types and actual instruments that measure general proficiency irrespective of the actual course followed. For this purpose, rating scales such as the ASLPR would seem most useful and should be used as the basis at least of the exit-point assessment.

III.4 Selecting Items

Advice on the activities used in applying rating scales is provided in the various papers published on the ASLPR and other scales [e.g., Ingram 1982, 1982a, 1984 and the Film Australia video kit and

associated package]. In assessing all four macroskills, essentially the same principles are involved, i.e., the aim is to elicit maximum language performance so that the learner's language behaviour can be observed and matched against the rating scale for each macroskill (speaking, listening, writing and reading) separately. At all times, it is the learner's practical skills that are to be assessed, i.e., knowledge mobilized in carrying out language tasks, not knowledge per se. This means that learners should be set real-life language tasks to perform with their proficiency in each macroskill being rated as their language behaviour is observed.

Guidance as to the suitability of activities is provided by the rating scales themselves since each behavioural description describes the sorts of activities in which learners can engage and, in the ASLPR, these are further identified through specific examples. As a general principle, it should be remembered that language ability is the ability to communicate meaning through organized words and practical proficiency can be observed only as the learner seeks to communicate (i.e., send or receive meaning). Real-life tasks have a purpose, a physical, social, and generally linguistic context, entail functions and, most fundamentally, involve an information gap to be bridged. Although tasks may often involve two or more macroskills, rating scales such as the ASLPR (though not all rating scales (2)) separate out the four macroskills and so, in interviewing, observing and rating, one has to focus on each macroskill and ensure that deficiencies in one do not intrude on the judgement made about another.

III.5 Relating Ratings and Grades through Goals

One of the difficulties teachers sometimes have in confronting rating scales is to know how to convert proficiency ratings to the sorts of grades traditionally used in Secondary Schools. Not least, if one is using the ASLPR, it may seem demotivating to state at the end of Year 10 that a learner has reached a proficiency of 1- in all four macroskills. Such concern ignores at least two important issues. First, there is no doubt that much of the disillusionment students feel with language study comes from false and unrealistic expectations. That is, on the one hand, they seem initially to believe that they will rapidly become like native speakers and they become disillusioned when they realize that language learning is a long and often arduous process and that they may never become native-like and, on the other hand, they believe (partly as a result of the exaggerated importance schools have traditionally attached to detailed formal accuracy) that they cannot communicate until they know and use the language with the accuracy (or, even, better than the accuracy) of native speakers.

The response to this disillusionment should take two related forms. First, if teaching at all times emphasizes communication, not only will it ensure that real, purposeful and meaning-oriented language and realistic language use are being taught, but the learner, on the one hand, will realize at all times that he can and does communicate even when his language forms are fragmentary and his range of vocabulary and other features limited and, on the other hand, there is invoked the same basic motivation (viz., the desire to communicate more

adequately) as occurs in natural first or second language learning. Second, syllabus writers and teachers should set clear and realistic goals, they should avoid seeming to promise the dishonest, "pie in the sky if you work hard enough to die" attainments that language syllabuses have traditionally implied. The Queensland German syllabus, for example, sets as the proficiency in Speaking for "Very High Achievement" in Year 10 as :

> Student demonstrates ability to ask and answer simple questions on the topics prescribed without hesitation. He can maintain a simple face-to-face conversation. He can speak with no significant errors in pronunciation, intonation and stress. Errors in grammar, syntax and vocabulary are few, and of a minor nature, and should not prevent comprehension by the listener. He shows a mastery of simple idiom and can use a variety of vocabulary within the prescribed topics. Self-confidence rather than self-consciousness is evident.

Such an expectation of non-native speakers in, at best, a 220 hour course spread over 3 years is entirely unrealistic and creates false expectations. Rating scales provide a ready means to set goals quite precisely. Thus, one might state that the goals in a three year unintensive French course (Years 8 to 10) is the proficiency S:1-, L:1-, W:1-, R:1- or, in Japanese, S:1-, L:1-, W:0+, R:0+.

This approach has several benefits. First, it is more likely to be honest and realistic. Second, it forces syllabus writers and teachers to think more clearly and realistically about appropriate course content and learning activities. Third, if the goals seem modest yet, in terms of what firm data is available on learning rates [e.g., FSI 1973], realistic then it can also be used to force policy-makers, education administrators, and curriculum planners to realize what the effect on language teaching and learning is when the allocated time is reduced to absurdly low levels either by starting foreign language study too late or by reducing the weekly allocation to 1½ or 2 hours. Fourth, goals identified in this way continually remind teachers of the essential differences referred to earlier between formal learning and acquisition, between what is formally presented and practised in class and what will be used in real language-use situations. This distinction leads to the differentiation between goals and objectives, the former being expressed in terms of the proficiency levels ultimately to be attained in the course, the latter identifying what the teacher needs to present in class and have the learner practise as the data unput to the learner's natural acquisition processes [cf. Ingram 1982; Section V]. Such a distinction is not only realistic in terms of the teacher's role as the source of the language in a foreign language context and in terms of the outcome of learning but it also has the effect in methodology of emphasizing the importance in the course of giving the learners many opportunities to use in realistic communication activities the language that is being formally presented.

Fifth, this approach to specifying goals readily lends itself to conversion into whatever grades (e.g., words, numbers, or letters) an

education system chooses to use. Thus, if goals are specified in terms of the proficiency levels the average student could be expected to achieve, then achievement above or below this leads to higher or lower grades. The present writer's recommendations, for example, to the Foreign Languages Advisory Committee of the Queensland Board of Secondary School Studies on the specification of goals for the unintensive Queensland Secondary School French courses and the conversion of proficiency levels attained to Board grades were as follows :

YEAR 10

GRADES	PROFICIENCY				
	S	L	R	W	OVERALL
Very High Achievement	>1	>1	>1	>1	All four macroskills >1
High Achievement	1	1	1	1	Two or more macroskills 1, others 1-
Sound Achievement	1-	1-	1-	1-	All four macroskills 1-
Low Achievement	0+	0+	0+	0+	1 or more macroskills 0+
Very Low Achievement	<0+	<0+	<0+	<0+	Two or more macroskills <0+

NOTE : Any variation below the number of macroskills at a specified level of achievement would have to be compensated for by an equal number above for the learner to be graded at that level of achievement.

YEAR 12

GRADES	PROFICIENCY				
	S	L	R	W	OVERALL
Very High Achievement	⩾2	⩾2	⩾2	⩾2	All four macroskills 2
High Achievement	2	2	2	2	Two or more macroskills 2, others 1+
Sound Achievement	1+	1+	1+	1+	All four macroskills 1+
Low Achievement	1	1	1	1	One or more macroskill 1
Very Low Achievement	<1	<1	<1	<1	Two or more macroskills 1

NOTE : Any variation below the number of macroskills at a specified level of achievement would have to be compensated for by an equal number above for the learner to be graded at that level of achievement.

IV Tests in Programme Planning

Traditionally, it has been held as a principle that assessment procedures should reflect teaching methods, objectives and goals. Nevertheless, as the present writer found in a survey of matriculation French examinations in Britain and Australia in the middle 1970's [Ingram 1978 : Chapter 2; Ingram 1977], tests (especially Public Examination papers) have too often become the syllabuses determining content and methods of courses. Though this is a serious misuse of tests, nevertheless tests of all sorts can have a wider input to language teaching programmes than merely occupying the last stage in the syllabus-teaching-learning-assessment sequence. We have already seen that developmental rating scales such as the ASLPR identify

learners' stages of development and indicate the path of future development and so they make a major contribution to programme planning by identifying developmental needs. Other tests (e.g., indirect or discrete point tests) that focus on specific linguistic items such as syntactic forms, vocabulary, discourse or function provide detailed information on what is or is not in the learners' linguistic competence and this can assist in identifying the detailed content to be included as the focus of teaching and, hence, the input to the acquisition process. In fact, all types of tests have a contribution to make to programme development though the nature of that contribution differs according to the nature of the test and how it is developed.

It is not possible to discuss in detail here the contribution of tests to programme development. Interested readers are referred to other papers where the issue has been discussed in some detail [see Ingram 1983; also 1982, 1984]. In the light of the concurrent development of interest in Australian language teaching in the development and measurement of "competency" or proficiency and in the use of graded objectives or graded levels of achievement, it is worth considering how rating scales that measure general proficiency relate to the sort of scales that comprise graded objectives or graded levels of achievement.

V Graded Objectives

The chief difference between a rating scale such as the ASLPR used to measure general proficiency and a set of graded objectives is that the former claims to be universally applicable within a concept of general proficiency, the non-specialist register, or everyday situations. Graded objectives constrict the proficiency sought to the ability to carry out specified tasks, in specified situations, generally using specified exponents or language elements. The two concepts can be in conflict if graded objectives are defined narrowly (as is their inherent risk) with the result that the objectives are achieved through the rote memorization of set phrases to carry out the specified tasks. In such an approach, students are tested by seeing whether they can repeat the set utterances in the specified situations. Rating scales designed to measure general proficiency are inappropriate to this approach because the notion of general proficiency and its development entail the hierarchical development of utterances and the natural (hence creative and not rote memorized) use of language.

However, the concepts of general proficiency and graded objectives (especially in Clark's sense of "graded levels of achievement "[e.g., Clark 1982, 1982a]) are not necessarily in conflict provided the graded levels of achievement have been designed to lead ultimately to the development of general proficiency. In fact, the ASLPR recognizes that formulaic (hence rote memorized) utterances are common at the 0+ level and even up to 1- and, to this extent, graded objectives even in the narrow sense may be relevant to the lowest proficiency levels. This seems to be implied in the first sentence of the proficiency

definition cited from the Queensland German Syllabus earlier but the rest of that definition misrepresents how language develops and suggests that near native-like "accuracy" can be imposed on utterances at any stage whereas, even with rote memorized utterances, this is unlikely in, for instance, "pronunciation, intonation and stress". From 1-, one expects to see more flexibility or creativity starting to emerge and, beyond this level, learners must be able to use creatively whatever language resources they have, i.e., to produce and understand original utterances. Graded levels of achievement courses beyond this level can be compatible with general proficiency but, to be so, the situations and tasks in which the language is to be taught and assessed are more clearly specified than in "general proficiency" courses, the situations, tasks and language exponents should be so selected that the learner's developmental needs are taken into account, and the language should be so experienced and practised that creativity is fostered, i.e., the learner is creating language in response to situations, purposes and meanings and not just repeating rote memorized utterances. Such graded objectives or graded levels of achievement courses lead to general proficiency. In such courses, both tests of general proficiency (hence, rating scales such as the ASLPR) and the more specific content-oriented tests (e.g., discrete point tests) are relevant, i.e., one would be interested in the level of general proficiency learners attain and use the ASLPR or another scale to measure it but one might also be interested in the learners' attainment on the graded objectives course content as well and use a variety of tests based directly on that content (tasks, functions, syntax, vocabulary, etc).

VI Conclusion

This paper has tried to consider the place of proficiency rating scales in the assessment of school language courses. While the present writer clearly believes that proficiency rating scales provide a valuable tool in language testing, it would be a serious mistake to believe they can displace all other test types. Tests of all sorts, indirect, semi-direct and direct, serve different purposes; direct tests such as rating scales are particularly useful for giving a meaningful measure of learners' practical language skills or practical proficiency but other tests serve other purposes equally well. Testing, like all language learning and teaching, should be seen as revolving around the "respond to need" principle: one must first ask why one is testing, for what purpose one is testing, what the constraints on the testing are, and then resolve on the most appropriate test to use. If the aim of the test is to measure general proficiency or if the aim of the course to be assessed is to develop general proficiency, then a direct instrument, a rating scale such as the ASLPR, would seem the most appropriate. If, however, the aim of the course is different (e.g., to develop knowledge) or the purpose of the testing is different (e.g., diagnosis), then some other test type may be more appropriate. In brief, the first step in testing is to identify what one wants to find out, to identfy what the purposes and possibilities of the available tests are, and then to select the test or tests that will provide the information required.

As we observed at the start of this paper, interest has grown rapidly in competency- or proficiency-based assessment especially as more emphasis has fallen on communicative goals and methods. In some places, consideration has been given to developing new rating scales for the new syllabuses emerging. Existing rating scales have their own characteristics, possibilities and limitations. However, the development of an instrument such as the ASLPR or any other reputable instrument is a long and arduous process. It requires considerable linguistic and psycholinguistic input, many hours of trialling, experimentation, and formulation and re-formulation of proficiency descriptions. The development of a scale requires the resolution of many complex theoretical and practical issues: the relation between sociocultural dialects, the relation between different macroskills, the relation between needs and universality of development, the apparent conflict between the situation-dependent nature of language and the concept of general proficiency, the relation between general proficiency and "special purpose" proficiency, and so on. Rather than develop new basic scales, it would seem preferable to devote resources to considering how existing scales can be related to particular courses in particular school contexts and to training teachers brought up in a more traditional testing context to understand and use appropriately the whole range of test types and instruments available. The emergence of communicative approaches to language teaching and the emergence of graded objectives have done much to make language teaching more practical in its goals and approach and more sensitive to the real learning and communication needs of both the community and the individual learners. The field of language testing has also changed and developed in response to similar issues, in particular providing, through instruments such as the ASLPR, a means to prepare more coherent and needs-related programmes, to specify practical goals, and to provide meaningful, readily interpretable statements of learners' practical language skills, i.e., of their competency or proficiency. It is now up to educational administrators, school syllabus designers, and teachers themselves to ensure that these innovations are used to improve the quality of the programmes we offer in Australian schools.

FOOTNOTES

(1) Some of the ideas in this paper were explored originally in a paper the author and others prepared for the Foreign Languages Advisory Committee of the Queensland Board of Secondary School Studies and entitled "Assessment: A Long-Term View".

(2) The FSI Scale, for instance, has only two sub-scales: Speaking (which includes Listening) and Reading.

REFERENCES

Clark, John L. 1982. "Graded Levels of Achievement in Foreign Language Learning (Lothian Region's Project in Modern Languages in Schools)". In Language Testing Newsletter, No. 2, Spring 1982 : 2-4.

Clark, John L. 1982a. "How are we to assess language learning in schools?" Paper to the Fourth National Languages Conference, Perth, 3rd-6th September, 1982. Reprinted in MacPherson 1982.

Foreign Service Institute School of Language Studies (FSI). 1973. "Expected levels of Absolute Speaking Proficiency in Languages Taught at the Foreign Service Institute". Washington, D.C. : U.S. Department of State. mimeograph.

Hyltenstam, Kenneth and Manfred Pienemann. 1984. Instructional and Social Implications of L2 Acquisition Research. London : Multilingual Matters, 1984.

Ingram, D.E. 1977. "Language Proficiency at Matriculation". In Australian Review of Applied Linguistics, Vol. 1, No. 1, December 1977 : 20-32.

Ingram, D.E. 1978. "An Applied Linguistic Study of Advanced Language Learning", thesis for the degree of Doctor of Philosopy, University of Essex, Colchester, England. In ERIC Collection ED168359.

Ingram, D.E. 1982. Introduction to the Australian Second Language Proficiency Ratings (ASLPR). Canberra: Department of Immigration and Ethnic Affairs/AGPS. 1979 rev. 1982.

Ingram, D.E. 1982a. "New Approaches to Language Testing". Paper to the Fourth National Languages Conference of the AFMLTA, Perth, 3rd-6th September, 1982. Reprinted in MacPherson 1982.

Ingram, D.E. 1983. "Needs, Tests and Language Programmes". Paper to the Eighth National Conference of the Australian Association of Special Education, 'Exceptional People : A Lifetime Commitment', Brisbane, 25th-29th June, 1983. Forthcoming in an RELC publication.

Ingram, D.E. 1984. "Assessing Proficiency: An Overview on some Aspects of Testing". Forthcoming in Hyltenstam and Pienemann, 1984.

Krashen, S.D. 1976. "Formal and Informal Linguistic Environments in Language Acquisition and Language Learning". In TESOL Quarterly, Vol. 10, No. 2, June 1976 : 157-168.

MacPherson, Anne. 1982. The Language Curriculum in the 1980's : Affective, Effective or Defective? Perth : MLTAWA/AFMTA.